Age and education adjusted normative data and discriminative validity for Rey's Auditory Verbal Learning Test in the elderly Greek population

Lambros Messinis, Grigorios Nasios, Antonios Mougias, Antonis Politis, Petros Zampakis, Eirini Tsiamaki, Sonia Malefaki, Phillipos Gourzis & Panagiotis Papathanasopoulos

To cite this article: Lambros Messinis, Grigorios Nasios, Antonios Mougias, Antonis Politis, Petros Zampakis, Eirini Tsiamaki, Sonia Malefaki, Phillipos Gourzis & Panagiotis Papathanasopoulos (2015): Age and education adjusted normative data and discriminative validity for Rey's Auditory Verbal Learning Test in the elderly Greek population, Journal of Clinical and Experimental Neuropsychology, DOI: 10.1080/13803395.2015.1085496

To link to this article: http://dx.doi.org/10.1080/13803395.2015.1085496

Published online: 20 Nov 2015.
Age and education adjusted normative data and discriminative validity for Rey’s Auditory Verbal Learning Test in the elderly Greek population

Lambros Messinis, Grigoris Nasios, Antonios Mougias, Antonis Politis, Petros Zampakis, Eirini Tsiamaki, Sonia Malefaki, Phillipos Gourzis and Panagiotis Papathanasopoulos

Neuropsychology Section, Department of Neurology, University of Patras Medical School, Patras, Greece; Higher Educational Institute of Epirus, Ioannina; Department of Speech and Language Therapy, Ioannina, Greece; Alzheimer Center of the Greek Psychogeriatric Association "Nestor," Athens, Greece; Department of Psychiatry, Eginition Hospital, School of Medicine, University of Athens, Athens, Greece; Department of Radiology, University of Patras Medical School, Patras, Greece; Department of Mechanical Engineering & Aeronautics, University of Patras (statistics), Patras, Greece; Department of Psychiatry, University of Patras Medical School, Patras, Greece

ABSTRACT
Rey’s Auditory Verbal Learning Test (RAVLT) is a widely used neuropsychological test to assess episodic memory. In the present study we sought to establish normative and discriminative validity data for the RAVLT in the elderly population using previously adapted learning lists for the Greek adult population. We administered the test to 258 cognitively healthy elderly participants, aged 60–89 years, and two patient groups (192 with amnestic mild cognitive impairment, aMCI, and 65 with Alzheimer’s disease, AD). From the statistical analyses, we found that age and education contributed significantly to most trials of the RAVLT, whereas the influence of gender was not significant. Younger elderly participants with higher education outperformed the older elderly with lower education levels. Moreover, both clinical groups performed significantly worse on most RAVLT trials and composite measures than matched cognitively healthy controls. Furthermore, the AD group performed more poorly than the aMCI group on most RAVLT variables. Receiver operating characteristic (ROC) analysis was used to examine the utility of the RAVLT trials to discriminate cognitively healthy controls from aMCI and AD patients. Area under the curve (AUC), an index of effect size, showed that most of the RAVLT measures (individual and composite) included in this study adequately differentiated between the performance of healthy elders and aMCI/AD patients. We also provide cutoff scores in discriminating cognitively healthy controls from aMCI and AD patients, based on the sensitivity and specificity of the prescribed scores. Moreover, we present age- and education-specific normative data for individual and composite scores for the Greek adapted RAVLT in elderly subjects aged between 60 and 89 years for use in clinical and research settings.

Neuropsychological tests are essential for a reliable and valid assessment of cognitive functioning and have been widely recognized for this purpose (Lezak, Howieson, Bigler, & Tranel, 2012). However, diagnostic decisions and rehabilitation planning can only be meaningful if culture-, language-, and demographic-specific normative data are available.

The lack of normative data for commonly used neuropsychological tests in Greece has led neuropsychologists working in Greece to either develop culture- and language-specific new tests (Folia & Kosmidis, 2003; Kosmidis, Vlahou, Panagiotaki, & Kiosseoglou, 2004; Vlahou et al., 2013) or collect normative data for commonly used neuropsychological tests developed in other countries (Aretouli & Kosmidis, 2006; Argirokastriotou, Samanda, & Messinis, 2005; Economou, 2003; Giannakou & Kosmidis, 2006; Konstantinopoulou et al., 2011; Messinis, Lada, et al., 2006; Messinis, Tsakona, & Papathanasopoulos, 2006; Papathanasiou, Messinis, Georgiou, & Papathanasopoulos, 2014; Zalonis et al., 2009).

In an effort to contribute to the development of appropriate culture-, language-, and demographic-specific normative data, we aimed to establish normative and discriminative validity data for the RAVLT in the elderly population using previously adapted learning lists for the Greek adult population. We administered the test to 258 cognitively healthy elderly participants, aged 60–89 years, and two patient groups (192 with amnestic mild cognitive impairment, aMCI, and 65 with Alzheimer’s disease, AD). From the statistical analyses, we found that age and education contributed significantly to most trials of the RAVLT, whereas the influence of gender was not significant. Younger elderly participants with higher education outperformed the older elderly with lower education levels. Moreover, both clinical groups performed significantly worse on most RAVLT trials and composite measures than matched cognitively healthy controls. Furthermore, the AD group performed more poorly than the aMCI group on most RAVLT variables. Receiver operating characteristic (ROC) analysis was used to examine the utility of the RAVLT trials to discriminate cognitively healthy controls from aMCI and AD patients. Area under the curve (AUC), an index of effect size, showed that most of the RAVLT measures (individual and composite) included in this study adequately differentiated between the performance of healthy elders and aMCI/AD patients. We also provide cutoff scores in discriminating cognitively healthy controls from aMCI and AD patients, based on the sensitivity and specificity of the prescribed scores. Moreover, we present age- and education-specific normative data for individual and composite scores for the Greek adapted RAVLT in elderly subjects aged between 60 and 89 years for use in clinical and research settings.
specific neuropsychological test norms for the Greek population, in 2007 we presented normative data for the adult population of a commonly used neuropsychological measure of verbal learning and episodic memory, the Rey Auditory Verbal Learning Test (RAVLT; Messinis, Tsakona, Malefaki, & Papatkanosoupolos, 2007; Rey, 1964). The RAVLT has been translated and normed in several languages and cultures (e.g., Ferreira Correira & Campagna Osorio, 2014; Lannoo & Vingerhoets, 1997; Lee, 2003; Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2005) with norms showing significant variance regarding their relationship with demographic variables.

Difficulty with memory and learning is one of the most frequent complaints of individuals with neuropsychological disorders and may contribute significantly to disability in activities of daily living and functional outcome (Lezak et al., 2012; Malec & Thompson, 1994; Schoenberg et al., 2006). Memory complaints in outpatient settings also appear to be the most frequent reason for neuropsychological referral (Lezak et al., 2012). Moreover, many common neurological and psychiatric disorders produce deficits in memory processes (Papathanasiou et al., 2014; Schoenberg et al., 2006; Vlahou et al., 2013).

Decline in episodic memory constitutes one of the core neuropsychological symptoms, particularly in elderly individuals at risk of developing Alzheimer’s disease or during the preclinical stages of dementia known as mild cognitive impairment (MCI; Lezak et al., 2012; Speer et al., 2014). Due to a significant increase in elderly patients’ subjective memory complaints (Zygouris & Tsolaki, 2015) and reports (Mitchell & Shiri-Feshki, 2009; Summers & Saunders, 2012) that patients with MCI have elevated rates of conversion to dementia, the objective and accurate assessment of episodic memory in this population is crucial for diagnostic purposes.

The RAVLT (Rey, 1964; Messinis et al., 2007; Schmidt, 1996) is a multitrial verbal learning test that is preferred to other tests under conditions of limited assessment time and in cases in which clinicians are solely interested in list-learning abilities and wish to dissociate these abilities from conceptual organization abilities such as those tapped by the California Verbal Learning Test (Lezak et al., 2012). The test affords the ability of a person to encode, consolidate, store, and retrieve verbal information and is usually applied using a five-trial presentation of a 15-noun word list (List A; Trials 1, 2, 3, 4, 5; presentation rate of one word per second), a single presentation of a 15-noun word interference list (List B; Trial 6), two postinterference recall trials (one immediate, Trial 7, and one delayed, Trial 8, D, ranging from 20 to 45 minutes, most commonly around 30 minutes), and a recognition trial (Trial 9, R) of 50 words containing the target words of Lists A and B and 20 distractor words phonetically or semantically similar to those in Lists A and B (Lezak et al., 2012; Schmidt, 1996).

Although several cognitive processes can be extracted by analyzing performance of the previously mentioned individual trials, it has been suggested that computing composite scores, which aggregate several trials, may provide a purer index of a specific cognitive process (Vakil, Greenstein, & Blachstein, 2010). Furthermore, composite scores may better reflect theoretical indices of memory than raw scores (Lezak et al., 2012). These scores are more informative than the scores for the single learning trials, as each score alone does not reflect learning (Lezak et al., 2012). More specifically, learning ability may be better reflected by two such composite measures/scores generated from the RAVLT. The first is Trial 5 minus Trial 1, which would reflect the learning rate. Performance on this RAVLT composite measure was found to be deficient in mild cognitive impairment (MCI) and Alzheimer’s disease (AD) patients (Nordlund et al., 2007). The second is the total score of all five learning trials, which reflects total acquisition/learning. Normative data are available for this measure from various studies covering a broad age range for both genders (Geffen, Moar, O’Hanlon, Clark, & Geffen, 1990; Schmidt, 1996; Vakil et al., 2010).

Another factor that may influence RAVLT performance is proactive and retroactive interference. Proactive interference takes place when previously learned material contributes negatively to acquisition or recall of new information. Trial B alone cannot reflect interference without being compared to Trial 1, which serves as a baseline. Older age increases susceptibility to proactive interference (Vakil et al., 2010). Retroactive interference occurs when subsequently presented material negatively influences the recall of previously learned material (Lezak et al., 2012; Vakil et al., 2010).
The long-term retention and forgetting rate may also be assessed by the RAVLT by testing the recall of List A (Trial 8) after a 20–45-minute delay period. However, Vakil and Blachstein (1997) suggest that viewing Trial 8 (D) as reflecting delayed recall or retention ability is insufficient if it is not compared to Trial 5, which serves as the baseline for the number of words learned. Thus, a more accurate measure of delay recall is Trial 5 minus Trial 8. Patients with MCI and Alzheimer’s disease who perform low on Trial 8 have significant deficits on long-term retention (Gunther, Holtkampa, Jolles, Herpetz-Dahlamanna, & Konrad, 2004). Similarly, for the recognition measure (Trial 9; retrieval efficiency), without a comparison with the preceding delayed recall trial (Trial 8), interpretation of the score on Trial 9 is problematic. Thus, it is recommended that the number of the words in Trial 8 is subtracted from that in Trial 9 in order to reflect retrieval efficiency (Vakil et al., 2010). Difficulties in retrieval are common in patients diagnosed with MCI (Broder, Herwig, Teipel, & Fast, 2008) and Alzheimer’s disease (Gainotti, Marra, & Villa, 2001). Early Alzheimer’s disease patients recall few words on Trial 1 and reach to about 6 words by Trial 5. They also have particular difficulty recalling words after a delay with distraction and recognize about two more words than they recall, with many intrusion errors (Lezak et al., 2012).

The RAVLT has shown adequate utility in discriminating cognitively healthy elderly from patients with MCI and preclinical AD (Estevez-Gonzalez, Kulisevsky, Boltes, Otermin, & Garcia-Sanchez, 2003; Siraly et al., 2015). Moreover, false-alarm responses in the RAVLT recognition trials, associated with more widespread temporal brain area function, may be a helpful marker for early diagnosis of MCI (Zeidman et al., 2008). The RAVLT has also shown adequate discriminatory ability when administered repeatedly in differentiating subjects with MCI from those who were initially diagnosed with MCI but subsequently reverted to normal (Antuono et al., 2007).

A large body of research shows that demographic variables, most notably age and years of formal education, and less consistently intelligence level and gender, influence performance on the RAVLT. The influence of age on RAVLT performance has been reported by numerous studies and has gained wide acceptance as the most influential demographic variable on RAVLT performance (Crossen & Wiens, 1994; Ferreira Correia & Campagna Osorio, 2014; Speer et al., 2014; Geffen et al., 1990; Messinis et al., 2007; Uchiyama et al., 1995; Van der Elst et al., 2005).

In the adult population, two segments can be distinguished: the younger adult age groups (20 to 59) years and the older adult age groups (60 to 90) years. For the younger adult segment, RAVLT performance for younger subjects is not as differentially sensitive as performance obtained by the older age groups, possibly suggesting different storage capacities and strategy utilization in the two adult segments (Lezak et al., 2012; Strauss, Sherman, & Spreen, 2006). Utilizing a Hebrew version of the AVLT, Vakil and Blachstein (1997) noted modest changes in AVLT performance in participants below the age of 60 compared to increasingly reduced recall in participants over 60 years old. Healthy elderly subjects, in comparison to younger ones, also tend to show greater forgetting rates of words at the end of the list during delayed recall, suggesting that older subjects rely more on short–lived memory processes—that is, immediate recall of the last words on the list—than do younger subjects (Lezak et al., 2012).

The influence of gender, years of formal education, and intellectual level appears less consistent across various studies (Lezak et al., 2012; Schmidt, 1996; Strauss et al., 2006). More specifically, Geffen et al. (1990) reported better performance in women adult participants than in men in the age range of 16–86 years. Lannoo and Vingerhoets (1997) noted that women, young adults, and persons with higher educational levels outperformed men, older adults, and individuals with fewer years of formal education. Generally, when studies report differences in performance due to influences of gender, women tend to outperform men on the recall trials, but not on the recognition trials (Geffen et al., 1990; Lannoo & Vingerhoets, 1997; Miatton, Wolters, Lannoo, & Vingerhoets, 2004; Speer et al., 2014; Strauss et al., 2006). Other studies, however (Harris, Ivnik, & Smith, 2002; Van der Elst et al., 2005), reported better performance of women on the recognition trials in addition to recall trials. The recognition trials measure how much was learned, regardless of the efficiency of spontaneous retrieval (Strauss et al., 2006). The recognition process allows for the distinction between problems with registration and storage from those of inefficient recall (Powell & Cripe, 1991). Normally, if the patient’s problem is simply
difficulty in retaining new information, then the recognition score will be little better than the delay recall trial score (Strauss et al., 2006). There are, however, reports in the literature of nonsignificant contributions of gender to RAVLT performance (Ferreira Correia & Campagna Osorio, 2014; Forrester & Geffen, 1991; Mitrushina, Boone, Razani, & D’Elia, 2005; Savage & Gouvier, 1992).

Regarding the specific contribution of intelligence levels to RAVLT performance, the trend is that recall is better in persons with higher intelligence levels (Strauss et al., 2006; Vakil & Blachstein, 1997). Steinberg, Bielauskas, Smith, Ivnik, and Malec (2005) reported the influence of intellectual functioning on RAVLT performance and presented age and intelligence level adjusted normative data for older adults (Mayo’s Older Americans Normative Studies, MOANS). Steinberg et al. (2005) further reported stronger correlations between RAVLT trials and full-scale IQ (FSIQ) at moderate levels of intelligence.

The literature regarding the specific influence of formal education on RAVLT scores has been contradictory. There are studies that have reported better performance in persons with higher educational levels (Lannoo & Vingerhoets, 1997; Miatton et al., 2004; Van der Elst et al., 2005), while others have not reported a significant contribution of education to RAVLT performance (Mitrushina et al., 2005; Mitrushina, Satz, Chervinsky, & D’Elia, 1991; Wiens, McMinn, & Crossen, 1988). In addition, certain authors noted that education does not account for RAVLT performance beyond that which is accounted for by intelligence level (Strauss et al., 2006).

As stated previously, in 2007, normative data (Messinis et al., 2007) were published for the RAVLT using newly adapted learning lists for the Greek adult population. One of the main limitations in that normative study was the lack of stratified subgroups for elderly participants over the age of 60. Elderly participants normally show a more distinguishable pattern of performance decline with advancing age, and it would have been preferable to have included narrower groupings. Another limitation was the unavailability of data for persons aged 80 and over.

In order to overcome these limitations, we assessed a relatively large group of healthy elderly adults and provide normative data based on demographic characteristics specific to the Greek culture and language for the RAVLT in this older population. We also examined the test’s validity in differentiating the performance of healthy elderly persons from that of patients with episodic memory deficits, by assessing large samples of patients diagnosed with amnestic mild cognitive impairment (aMCI) and Alzheimer’s disease (AD) and comparing their performance to an age-, gender-, and education-matched cognitive healthy control group. Moreover, we explored the diagnostic utility of the test through receiver operator characteristic (ROC) analyses in patients with aMCI and AD and provide cutoff scores, stratified by age, in discriminating cognitively healthy controls from aMCI and AD patients, based on the sensitivity and specificity of the prescribed scores.

Method

Participants

Two hundred and fifty eight (258) cognitively healthy elderly native Greek speakers (161 females or 62.4%) recruited primarily from southwestern Greece and Athens, took part in the present study voluntarily and after providing written consent for their participation. Healthy participants were contacted/approached and if interested were invited to take part in the study by the staff (neurologists, psychiatrists, psychologists, and neuropsychologists) of the outpatient Neuropsychology Unit, Department of Neurology, University of Patras Medical School and the Alzheimer Center of the Greek Psychogeriatric Association “Nestor,” in Athens, Greece. These participants were approached mainly through senior citizen centers and our outpatient clinics, as stated above.

Healthy participants were aged between 60 and 89 years (age: $M = 74.21$ years, $SD = 7.71$; level of education: $M = 9.85$ years, $SD = 4.59$). Exclusion criteria for the healthy participants were a history of psychiatric, neurological, or cardiovascular disorders or of substance abuse or dependence (including alcohol and benzodiazepine abuse), any other medical condition (including hearing impairment) that might affect neuropsychological performance, and non-native speakers of the Greek language. We further excluded from the study elderly adults who on initial testing obtained scores of less than 27 on the Greek validated version of the Mini-Mental State Examination (Fountoulakis, Tsolaki, Chantzi, & Kazis, 2000), a
brief screening measure for global cognitive deficits. Elderly adults who met a current Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM–5; American Psychiatric Association, 2013), diagnosis of major neurocognitive disorder (MND) due to Alzheimer’s disease. AD patients had a mean age of 75.55 years (SD = 6.38) and a mean level of education of 10.60 years (SD = 4.24).

Procedure
Healthy participants were tested individually by psychologists at either of the previously mentioned clinical settings. These participants were initially screened through a standardized interview at the beginning of the testing session by the project staff clinical neuropsychologist and physician (neurologist or psychiatrist), in order to exclude those with health problems or other exclusion criteria as described above. They were also administered the Greek validated MMSE (Fountoulakis et al., 2000) and the Greek validated version of the Geriatric Depression Scale (Fountoulakis et al., 1999). Psychologists taking part in the testing of participants had been intensively trained in the administration procedures of various neuropsychological measures, including the RAVLT, by doctoral-level clinical neuropsychologists. Participants were then assessed using Greek adapted lists of the RAVLT (Messinis et al., 2007). These adapted lists were based partly on the original English lists presented in Schmidt (1996, p. 71; List A, learning list, and List B, interference list) and Lezak, Howieson, and Loring (2004, p. 423). The original English words presented in the above lists were initially translated without change to the Greek language. From these two original English word lists we retained 10 of the original words from List A and seven of the original words from List B (interference trial) as presented in Schmidt (1996, p. 71). The remaining words used for the development of the Greek lists were new words adapted for the Greek language. The word items on these newly formed Greek lists were evaluated for consistency on the following dimensions: All words were two- or three-syllable concrete nouns; there were no obvious semantic or phonetic associations or similarities between words in the same list; all were common words that are normally acquired in Greek-speaking persons with relatively low levels of education; and all words had frequent occurrence in the Greek language. The probability of the occurrence of the word in common usage in the Greek language was
ascertained using the Institute for Language and Speech Processing Greek Corpus (Hatzigeorgiu et al., 2000). Using the above criteria, form equivalence was established between the two new lists. These two new lists yielded comparable mean scores for the various trials of the RAVLT, as there were no significant differences between the two lists on the matching consistency dimensions, as described previously (Messinis, Tsakona, et al., 2006). The original Lists A and B (Lezak et al., 2004, p. 423; Schmidt, 1996, p. 71) and the two new Greek lists (A1 and B1) in the order that the words were read to participants are presented in Table 1. The administration procedure used was the one originally followed by Geffen et al. (1990), as referred to in Schmidt (1996) and Lezak et al. (2004). We must note, however, that after participants completed Trial B (interference trial), we administered other nonverbal neuropsychological tests (Symbol Digit Modalities Test; Benton Visual Retention Test; Color Trails Test Parts 1 and 2) for approximately 25 minutes in order to avoid interference with previously learned verbal material (Argirokastritou et al., 2005; Messinis, Lyros, Georgiou, & Papathanasopoulos, 2009, Messinis, Malegiannaki, Christodoulou, & Papathanasopoulos, 2011). After this 25-min delay period, we asked participants to recall the words learned in trials 1–5 (delay recall trial) and then administered the recognition trial asking the patients to identify the target words of List A.

Patients diagnosed with aMCI or AD were also tested individually by psychologists at either of the previously mentioned clinical settings. They had been previously diagnosed by a multidisciplinary consensus group composed of neurologists, psychiatrists, neuropsychologists, and speech pathologists utilizing the criteria mentioned above. They were initially administered the Greek validated MMSE (Fountoulakis et al., 2000) and the Greek validated version of the Geriatric Depression Scale (Fountoulakis et al., 1999). They were then administered a brief neuropsychological battery comprising the RAVLT (Messinis et al., 2007), tests of executive functioning and psychomotor speed, including the color trails test (only Part 1; Messinis et al., 2011), the Symbol Digit Modalities Test (Argirokastritou et al., 2005), a test of visuospatial ability, the Clock Drawing Test (Bozikas, Giazkoulidou, Hatzigeorgiadou, Karavatos, & Kosmidis, 2008), and a test assessing nonverbal memory and visuocstruction, the Benton Visual Retention Test (Messinis et al., 2009). These two patient groups underwent RAVLT testing by the same administration procedure as that described previously for the healthy elderly group.

### Statistical analyses

We initially examined our data visually to determine whether the distributions met normality requirements. All data points that were considered outliers or extreme outliers were excluded from the analyses. Then, we calculated the main descriptive statistics for all studied variables and tested groups separately. The normality assumption of our data was also tested using the Kolmogorov–Smirnov test for normality. Generally, the normality hypothesis was rejected in almost all cases so the nonparametric Kruskal–Wallis test was used to examine whether the samples of the tested groups originate from the same distribution (healthy participants, aMCI patients, and AD patients). Post hoc pairwise comparisons between the tested groups were also provided. Furthermore, in order to examine the potential contribution of the demographic variables age, gender, and years of formal education on the performance of the different trials of the RAVLT in the full sample of healthy elderly adults, nonparametric Spearman partial correlation coefficients were computed, and their statistical significance was also tested. Finally, ROC

### Table 1. The original RAVLT (Lists A and B) and greek lists (A1 and B1).

<table>
<thead>
<tr>
<th>Original list</th>
<th>List A&lt;sup&gt;a&lt;/sup&gt;</th>
<th>List A&lt;sup&gt;b,c&lt;/sup&gt;</th>
<th>Greek lists</th>
<th>List B&lt;sup&gt;a&lt;/sup&gt;</th>
<th>List B&lt;sup&gt;b,c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>List A&lt;sup&gt;a&lt;/sup&gt;</td>
<td>List A&lt;sup&gt;b,c&lt;/sup&gt;</td>
<td>List A&lt;sup&gt;b,#&lt;/sup&gt;</td>
<td>List B&lt;sup&gt;b,#&lt;/sup&gt;</td>
<td>List B&lt;sup&gt;b,#&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Drum</td>
<td>Desk</td>
<td>Day</td>
<td>Mera</td>
<td>Desk</td>
<td>Thranio</td>
</tr>
<tr>
<td>Curtain</td>
<td>Ranger</td>
<td>Curtain</td>
<td>Kourtina</td>
<td>Number</td>
<td>Anthimos</td>
</tr>
<tr>
<td>Bell</td>
<td>Bird</td>
<td>Train</td>
<td>Treno</td>
<td>Bird</td>
<td>Pouli</td>
</tr>
<tr>
<td>Coffee</td>
<td>Shoe</td>
<td>Coffee</td>
<td>Kafes</td>
<td>Shoe</td>
<td>Papoutsi</td>
</tr>
<tr>
<td>School</td>
<td>Stove</td>
<td>School</td>
<td>Skolio</td>
<td>Child</td>
<td>Pedi</td>
</tr>
<tr>
<td>Parent</td>
<td>Mountain</td>
<td>Parent</td>
<td>Yoneas</td>
<td>Mountain</td>
<td>Vouno</td>
</tr>
<tr>
<td>Moon</td>
<td>Glasses</td>
<td>Hand</td>
<td>Heri</td>
<td>Water</td>
<td>Nero</td>
</tr>
<tr>
<td>Garden Towel</td>
<td>Garden</td>
<td>Kipos</td>
<td>Book</td>
<td>Vivilo</td>
<td></td>
</tr>
<tr>
<td>Hat</td>
<td>Cloud</td>
<td>Hat</td>
<td>Kapelo</td>
<td>Cloud</td>
<td>Sinefo</td>
</tr>
<tr>
<td>Farmer</td>
<td>Boat</td>
<td>Farmer</td>
<td>Agrotis</td>
<td>Boat</td>
<td>Varka</td>
</tr>
<tr>
<td>Nose</td>
<td>Lamb</td>
<td>Nose</td>
<td>Miti</td>
<td>Bridge</td>
<td>Yefira</td>
</tr>
<tr>
<td>Turkey</td>
<td>Gun</td>
<td>Street</td>
<td>Dromos</td>
<td>Woman</td>
<td>Yineka</td>
</tr>
<tr>
<td>Color</td>
<td>Pencil</td>
<td>Color</td>
<td>Hroma</td>
<td>Pencil</td>
<td>Molivi</td>
</tr>
<tr>
<td>House Church</td>
<td>House</td>
<td>Sphi</td>
<td>Button</td>
<td>Koubi</td>
<td></td>
</tr>
<tr>
<td>River Fish</td>
<td>Door</td>
<td>Porta</td>
<td>Body</td>
<td>Soma</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Original list (Lezak et al., 2004; Rey, 1964; Schmidt, 1996).
<sup>b</sup>Greek lists (Messinis et al., 2007).
<sup>c</sup>Order in which the Greek list was read to subjects and at a presentation rate of one word per second.
Correlations among RAVLT variables in healthy elderly.

Results

Initially, in order to examine thoroughly the shape of the studied variables, we created boxplots for all the variables for each patient group. As noted in Figure 1, there is no common pattern in the shapes of the distributions of the studied variables. In most of the cases, there is a discrepancy from normality, and also the distributions are skewed either positively or negatively. However, no floor or ceiling effect is present, with the exception of the RAVLT delay recall trial in the group with AD who show a floor effect. Moreover, the nonparametric Spearman correlation coefficients of RAVLT trials (variables) showed most items to be significantly correlated with each other (Table 2).

Effect of demographic variables on RAVLT test performance

In order to examine the potential contribution of demographic variables age, gender, and years of formal education on the performance of the different trials of the RAVLT in the full sample of healthy elderly adults, nonparametric Spearman partial correlation coefficients were computed, and their statistical significance was tested (see Table 3). In general, age and education significantly influenced the performance on almost all RAVLT variables examined. The older elderly with a lower education level performed worse than the younger elderly with higher education levels. Moreover, gender did not significantly influence performance on most of the RAVLT variables, with the exception of the delay recall and retention trials.

Age and education stratified normative data

Given the significant influence of education and age on performance to most of the RAVLT trials revealed by the analyses, and in order to generate normative data for the Greek elderly population over the age of 60, we grouped our sample into three age groups—60–69, 70–79, and 80–89 years —and also stratified our sample based on the level of education, so as to reflect actual school requirements in Greece (compulsory education is nine years): 1–9 years and above 9 years. In Table 4 we present normative data for the Greek healthy

Table 3. Contributions of education, age, and gender to RAVLT individual trial and composite scores.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Age</th>
<th>Education</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAVLT 1</td>
<td>.399*</td>
<td>.244*</td>
<td>.0431</td>
</tr>
<tr>
<td>RAVLT 5</td>
<td>.506*</td>
<td>.273*</td>
<td>.0092</td>
</tr>
<tr>
<td>RAVLT B</td>
<td>.4961*</td>
<td>.3615*</td>
<td>.1317</td>
</tr>
<tr>
<td>RAVLT 7</td>
<td>.2581*</td>
<td>.3449*</td>
<td>.1791</td>
</tr>
<tr>
<td>RAVLT 8 Rec</td>
<td>.2434*</td>
<td>.3541*</td>
<td>.1971</td>
</tr>
<tr>
<td>Learning rate</td>
<td>.2394*</td>
<td>.1414*</td>
<td>.0635</td>
</tr>
<tr>
<td>Retention</td>
<td>.2485</td>
<td>.0920</td>
<td>.2215</td>
</tr>
<tr>
<td>Retrieval</td>
<td>.0210</td>
<td>.0599</td>
<td>.0059</td>
</tr>
<tr>
<td>Total learning</td>
<td>.4681*</td>
<td>.3378*</td>
<td>.0620</td>
</tr>
</tbody>
</table>

Note. RAVLT = Rey Auditory Verbal Learning Test; RAVLT Trial 1; RAVLT 5 = RAVLT Trial 5; RAVLT B = RAVLT Trial B (interference trial); RAVLT 7 = RAVLT Trial 7 (immediate recall); RAVLT D = RAVLT Trial D (delay recall trial); RAVLT R = RAVLT Trial R (recognition trial); total learning = Σ Trials 1–5; learning rate = Trial 5 – Trial 1; retention = Trial 5 – Trial D; retrieval efficiency = Trial R – Trial D.

*p < .05.

Table 2. Correlations among RAVLT variables in healthy elderly.

<table>
<thead>
<tr>
<th>Variable</th>
<th>RAVLT 1</th>
<th>RAVLT 5</th>
<th>RAVLT B</th>
<th>RAVLT D</th>
<th>RAVLT R</th>
<th>Total learning</th>
<th>Learning rate</th>
<th>Retention</th>
<th>Retrieval</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAVLT 1</td>
<td>.571*</td>
<td>.547*</td>
<td>.512*</td>
<td>.418*</td>
<td>.782*</td>
<td>-.040</td>
<td>.081</td>
<td>.015</td>
<td></td>
</tr>
<tr>
<td>RAVLT 5</td>
<td></td>
<td>.416*</td>
<td>.714*</td>
<td>.522*</td>
<td>.882*</td>
<td>.775*</td>
<td>.194*</td>
<td>.042</td>
<td></td>
</tr>
<tr>
<td>RAVLT B</td>
<td></td>
<td></td>
<td>.335*</td>
<td>.522*</td>
<td>.093</td>
<td>.218*</td>
<td>.010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAVLT D</td>
<td></td>
<td></td>
<td></td>
<td>.607*</td>
<td>.774*</td>
<td>.519*</td>
<td>.426*</td>
<td>.346*</td>
<td></td>
</tr>
<tr>
<td>RAVLT R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.604*</td>
<td>.343*</td>
<td>-.105</td>
<td>.460*</td>
<td></td>
</tr>
<tr>
<td>Total learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-.499*</td>
<td>.092</td>
<td>-.036</td>
<td></td>
</tr>
<tr>
<td>Learning Rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-.168*</td>
<td>-.090</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-.214*</td>
<td>.428*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. RAVLT = Rey Auditory Verbal Learning Test; RAVLT 1 = RAVLT Trial 1; RAVLT 5 = RAVLT Trial 5; RAVLT B = RAVLT Trial B (interference trial); RAVLT D = RAVLT Trial D (delay recall trial); RAVLT R = RAVLT Trial R (recognition trial); total learning = Σ Trials 1–5; learning rate = Trial 5 – Trial 1; retention = Trial 5 – Trial D; retrieval efficiency = Trial R – Trial D.

*p < .01.
elderly population stratified by age and education level.

**Group differences in RAVLT performance**

In order to determine the validity of the RAVLT in discriminating the performance of the patient group from that of healthy elderly participants, we conducted comparisons on the RAVLT outcome parameters between the three groups (aMCI, AD patients, and cognitive healthy participants). The three groups were matched on gender ratio, $\chi^2(2) = 5.653, p = .059$, level of education, $F(2, 503) = 1.024, p = .360$, and age, $F(2, 503) = 1.615, p = .200$. Utilizing the Kruskal–Wallis non-parametric test, we noted a significant main group effect on Trial 1, $\chi^2(2) = 37.445, p < .001$, trial 5, $\chi^2(2) = 67.465, p < .001$, trial B, $\chi^2(2) = 25.707, p < .001$, trial D, $\chi^2(2) = 133.529, p < .001$, and trial R, $\chi^2(2) = 58.148, p < .001$. Moreover, we noted a significant main group effect for the composites total learning ($\sum$ Trials 1–5), $\chi^2(2) = 69.48, p < .001$, learning rate (Trial 5 – Trial 1), $\chi^2(2) = 41.3, p < .001$, and retention (Trial 5 – Trial D), $\chi^2(2) = 48.28, p < .001$.

Descriptive statistics for each trial and composite of the RAVLT are presented in Table 5, including pairwise comparisons between aMCI and healthy elderly, between AD and healthy elderly, and between aMCI and AD patients. These pairwise comparisons revealed significant differences between aMCI patients and the healthy elderly on most trials of the RAVLT, with the exception of Trial B (interference trial) and the composite learning rate. AD patients differed significantly from the healthy elderly on all the RAVLT trials examined. Further, comparison of the two clinical groups showed that the performance of AD patients was significantly different from that of aMCI patients on all individual trials. However, nonsignificant differences between the two clinical groups were noted on the composites retention and retrieval efficiency.

**Clinical utility**

We explored the clinical utility of the test through a receiver operator characteristic (ROC) curve in order to calculate sensitivity and specificity. Specifically, we examined the utility of the RAVLT individual trials and composites to discriminate cognitively healthy controls from aMCI
and AD patients. The area under the ROC curve (AUC), an index of effect size and overall diagnostic accuracy, showed that in most cases the RAVLT trials adequately differentiated between the performance of healthy elders and aMCI/AD patients, although there were some exceptions that are discussed below. We also provide cutoff scores in discriminating cognitively healthy controls from aMCI and AD patients, based on the sensitivity and specificity of the prescribed score (see Table 6).

### Discussion

Neuropsychological assessment is essential for the diagnosis of memory and other cognitive deficits in the elderly and for their effective care and treatment. Given the significant increase in the number of elderly in Greece, which has one of the highest aging rates in Europe (National Statistical Office of Greece, 2014) the need for standardized neuropsychological assessment tools that deliver high-quality information and are specific for culture, language, and demographic variables is urgent.

Despite the widespread use of neuropsychological measures in clinical and research settings in Greece in the last few years, normative data for common neuropsychological assessment measures especially for the elderly population remain largely unavailable. In an effort to contribute towards filling this gap, we generated culture- and language-specific normative data for the Greek elderly population of a widely used, easily administered measure of verbal learning and episodic memory, the RAVLT (Rey, 1964; Schmidt, 1996), adequately stratified by demographic variables that contributed significantly to RAVLT performance on verbal learning lists adapted for the Greek adult population (Messinis et al., 2007). We further provide data on the test’s clinical utility in discriminating elderly patients with verbal learning and episodic memory deficits. In our previous presentation of normative data for the adult population in Greece (Messinis et al., 2007), we did not adequately stratify our norms for elderly participants over the age of 60, and data for adults aged 80 and over were not available. To our knowledge there have been no other attempts to date to collect normative data for the RAVLT in the elderly population in Greece.

The RAVLT has gained wide recognition in the aging literature and has been used extensively for the assessment of multiple memory indices and memory processes, such as initial recall, consolidation, retrieval, recognition, and protractive and retroactive interference (Constantinidou et al., 2014; Speer et al., 2014). In general, older adults over the age of 60 perform worse than their younger counterparts on the learning trials of the RAVLT (Messinis et al., 2007; Constantinidou et al., 2014). Further, older adults diagnosed with amnestic mild cognitive impairment (aMCI), a transitional state between healthy aging and dementia, perform in the impaired range on this and other similar episodic verbal memory tests (Constantinidou et al., 2014; Estevez-Gonzalez et al., 2003).

The results of the present study revealed that age accounted for a substantial proportion of the variance in RAVLT performance favoring younger elderly participants—that is, a decline in performance was observed with increasing age in a linear fashion. This finding is consistent with reports in

---

Table 5. Group comparisons on the RAVLT individual trials and composites.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Healthy elderly (N = 258)</th>
<th>aMCI patients (N = 192)</th>
<th>AD patients (N = 65)</th>
<th>aMCI/healthy elderly p</th>
<th>AD/healthy elderly p</th>
<th>aMCI/AD p</th>
</tr>
</thead>
<tbody>
<tr>
<td>RAVLT 1</td>
<td>4.24</td>
<td>1.76</td>
<td>3.54 1.53</td>
<td>2.80 1.59</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>RAVLT 5</td>
<td>8.38</td>
<td>2.83</td>
<td>7.39 2.66</td>
<td>5.15 2.31</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>RAVLT B</td>
<td>4.06</td>
<td>2.01</td>
<td>3.74 1.29</td>
<td>2.62 1.92</td>
<td>.171</td>
<td>.000</td>
</tr>
<tr>
<td>RAVLT D</td>
<td>6.13</td>
<td>3.29</td>
<td>3.82 2.10</td>
<td>1.03 1.07</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>RAVLT R</td>
<td>10.51</td>
<td>3.38</td>
<td>8.70 4.18</td>
<td>6.09 3.30</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Total learning</td>
<td>33.12</td>
<td>10.15</td>
<td>28.57 9.25</td>
<td>21.07 9.08</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Learning rate</td>
<td>4.13</td>
<td>2.37</td>
<td>3.81 2.22</td>
<td>2.35 1.76</td>
<td>.401</td>
<td>.000</td>
</tr>
<tr>
<td>Retention</td>
<td>2.25</td>
<td>2.13</td>
<td>3.58 2.15</td>
<td>4.12 1.89</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Retrieval</td>
<td>4.43</td>
<td>2.89</td>
<td>4.91 3.08</td>
<td>5.06 3.93</td>
<td>.000</td>
<td>.000</td>
</tr>
</tbody>
</table>

Note. RAVLT = Rey Auditory Verbal Learning Test; aMCI = amnestic mild cognitive impairment patients; AD = Alzheimer’s disease patients; RAVLT 1 = RAVLT Trial 1; RAVLT 5 = RAVLT Trial 5; RAVLT B = RAVLT Trial B (interference trial); RAVLT D = RAVLT Trial D (delay recall trial); RAVLT R = RAVLT Trial R (recognition trial); total learning = Σ Trials 1–5; learning rate = Trial 5 – Trial 1; retention = Trial 5 – Trial D; retrieval efficiency = Trial R – Trial D.

Asymptotic significance (two-sided tests); the significance level is .05.
the literature that age is the most influential demographic characteristic on RAVLT performance (Crossen & Wiens, 1994; Geffen et al., 1990; Uchiyama et al., 1995; Van der Elst et al., 2005). The finding that age contributes significantly to RAVLT performance is also true for the elderly.
population, with past (Steinberg et al., 2005) and more recent (Speer et al., 2014) studies providing such evidence.

This age-dependent decline in episodic memory has been linked to various brain alterations, including decreased information processing speed due to reductions in brain volume (Raz, 2000). Further, white matter loss, especially in the prefrontal areas, is associated with deficits in tasks assessing attention, working memory, and executive functions, implying that aging leads to a natural loss of cognitive functions (Speer et al., 2014).

Reports regarding the effects of formal education on RAVLT performance have been mixed in the literature. One possible explanation for this discrepancy is that it depends on the amount of difference or the actual levels of education being considered. In this study, we found that level of education contributed significantly to most of the trials and composite scores on the RAVLT. Participants with lower levels of education in the formal Greek educational system (under 9 years of education) performed worse than participants with higher levels of education. Unlike most studies, as reported by Schmidt (1996), which have found relatively weak relationships between education and RAVLT performance, and suggest that education is not a substantial demographic characteristic contributing to this test, in the present study we showed that education is a substantial demographic contributor when making normative comparisons in older Greek adults over the age of 60. The present data, as regards the influence of education on RAVLT performance, support our previous study that included younger and older Greek adults (Messinis et al., 2007) and numerous other studies that have found similar results (Constantinidou et al., 2014; Lannoo & Vingerhoets, 1997; Miatton et al., 2004; Van der Elst et al., 2005) and that included older participants up to 86 years old. There are, however, studies that support the notion of Schmidt (1996) of relatively inconsistent and insignificant contributions of education to RAVLT performance (Mitrushina et al., 2005; Mitrushina et al., 1991; Wiens et al., 1988), including a recent study with older adults (Speer et al., 2014).

Figure 1. Boxplots of the studied variables for the elderly healthy, amnestic mild cognitive impairment patients (MCI), and Alzheimer’s disease (AD) patients. RAVLT = Rey Auditory Verbal Learning Test. To view this figure in color, please visit the online issue of the Journal.
In contrast to our previous study (Messinis et al., 2007) that revealed a significant contribution of gender to RAVLT performance in a sample of younger and older adults, in this study only the recognition trial and retention composite was influenced by gender. This is in keeping with the majority of the cited literature where the contribution of gender has appeared less consistent (Schmidt, 1996; Steinberg et al., 2005). When an effect of gender has been noted, the general trend is that women outperform men on the recall trials, but not on the recognition trials (Geffen et al., 1990; Lannoo & Vingerhoets, 1997; Miatton et al., 2004; Strauss et al., 2006), although some studies (Harris et al., 2002; Van der Elst et al., 2005) reported better performance of women participants on the recognition trials in addition to recall trials.

One possible reason that may explain the lack of gender dominance favoring females in RAVLT performance in our elderly sample in the present study compared to our previous study of younger and older adults (Messinis et al., 2007), which found such an effect, is that aging females have a reduced neuroprotective influence of female sex hormones (Speer et al., 2014). Moreover, aging may possibly reduce the known higher verbal ability of women, which has been linked to better performance in verbal memory tests (Lezak et al., 2012; Speer et al., 2014).

Regarding the influence of intelligence level to RAVLT performance in the present study, we chose not to examine the contribution of intelligence in our elderly Greek sample for two main reasons: First, when the study was initiated there were no available standardized tests of intelligence in Greece for adults/elderly. Secondly, if our data were found to be influenced by intelligence level, and thus stratified by this variable, this would require that participants tested for verbal learning and memory deficits would have to also complete an intelligence test to establish full-scale IQ, before the norms could be used adequately. In clinical outpatient settings where most of the elderly Greek with memory problems are assessed, intelligence testing, which is highly time consuming, requires specialized training to interpret, and is difficult to complete, especially in the elderly and the demented, would not allow for norms stratified by intelligence to be used.

As noted previously, a valid assessment of memory function is often essential for diagnosing memory decline. On clinical grounds, the diagnosis of preclinical dementia stages relies on the diagnosis of mild cognitive impairment (MCI). The criteria for this diagnosis subsume the presence of subjective memory complaints, especially the amnestic MCI subtype, preferably corroborated by a close informant, and documented by abnormal performance on cognitive testing, based on demographically adjusted normative data (Petersen, Doody, et al., 2001; Summers & Saunders, 2012). In this study, we compared the performance of elderly healthy elderly from patients with aMCI and Alzheimer’s dementia. We found that all trials of the test adequately differentiated the performance of aMCI patients from that of the healthy elderly, with the exception of the interference trial and the composite measure learning rate. Our findings regarding the ability of most RAVLT trials to differentially diagnose cognitively healthy elderly from aMCI patients support previous studies (Estevez-Gonzalez et al., 2003; Zeidman et al., 2008) that found similar results. However, although the finding regarding the learning rate may appear unexpected, Broder et al. (2008) note that both the elderly and MCI patients may show a slower rate of learning and improvement in tasks of memory and repeated learning. In contrast, all trials and composite measures differentiated the performance of AD patients from that of the healthy elderly. Various studies have noted that the RAVLT may help to identify patients with Alzheimer’s dementia and even distinguish with a high degree of accuracy between types of dementia (Estevez-Gonzalez et al., 2003; Ricci, Graef, Blundo, & Miller, 2012; Tierney et al., 1994).

When comparing the two clinical groups, we found that all individual trials differentiated the performance of AD patients from that of the MCI group. However, the composite measures retention and retrieval efficiency were found to be nonsignificant between the two clinical groups. This finding implies that both clinical groups show relatively increased rates of forgetting and retrieval efficiency and support the findings of Broder et al. (2008) and Gainotti et al. (2001), who stipulate that these deficits are common in both MCI and AD patients.

Regarding the results of the ROC analyses, we found that the trials with the highest overall diagnostic accuracy (AUC), independent of cutoff scores, for all age categories were the delay recall.
trial and the total learning composite measure that discriminated healthy elderly from AD patients. Overall, AUCs ranged from poor (i.e., age group 60–69 retention and retrieval composites) to very high with excellent discriminatory utility (i.e., 60–69 Trial 5 and Trial D). A large range of sensitivities, specificities, positive predictive values (PPVs), and negative predictive values (NPVs) were also recorded, ranging from low to high, for the various cutoff scores provided. In some cases the composite scores that were calculated assisted in formulating a more refined differential diagnosis, although very low AUCs and sensitivities, specificities, NPVs, and PPVs were also noted. In this respect, it is suggested that the clinician utilize the combination of these values in determining the accuracy of the trial (see Table 6).

An important question arising from our findings is what makes performance on the RAVLT sensitive to early decline in aMCI and AD patients. One possible explanation is that medial temporal lobe atrophy (MTA; hippocampal–parahippocampal formation) seems to be an important anatomical feature of AD and its prodromal stage, amnestic mild cognitive impairment (Peruzza Marchianni, Figuerdo Balthazar, Cendes, & Pereira Damasceno, 2008). Episodic memory—that is, the ability to acquire explicit information—is affected most in aMCI and AD. In contrast, memory related to events that occurred earlier in life (past events) remains intact for long periods of time. This pattern of impairment reflects deficits mainly in memory processes of encoding and storage, in contrast to retrieval, which remains intact until the late stages of AD (Peruzza Marchianni et al., 2008; Ricci et al., 2012). A recent study that investigated the relationship of the functional memory processes and medial temporal lobe atrophy in patients with aMCI and AD (Boon, Melis, Olde Rikkert, & Kessels, 2011) found significant associations with the encoding and storage processes, but not with retrieval. This finding implies that aMCI and AD patients show a decline in encoding and storage of verbal information that is related to MTA atrophy. Further, Peruzza Marchianni et al. (2008) noted significant correlations between hippocampal volumes and scores on the RAVLT, confirming that medial temporal lobe structures are closely related with verbal memory performance in normal aging as well as in aMCI and AD patients. It therefore appears that atrophy of the medial temporal lobe regions may predict the presence of future AD, as hippocampal degeneration appears to occur before the onset of overt dementia (Estevez-Gonzalez et al., 2003).

In evaluating the generalizability of our findings, several potential limitations need to be highlighted. First, we realize that by not generating intelligence adjusted normative data we may limit the interpretive accuracy of the RAVLT in this population. On the other hand, if such intelligence adjusted normative data were utilized in Greek settings, the test norms would most probably be ignored in the absence of intelligence testing availability. Secondly, decreased familiarity with neuropsychological assessment procedures, which differ from traditional medical procedures to which elderly individuals in Greece have become accustomed, may have also influenced our findings. Examiners were, however, well trained in the administration of the RAVLT and had previous experience with elderly participants. Significant efforts were made in order to ensure that these older participants correctly understood all administration procedures, therefore minimizing this possible limitation. Another potential limitation concerns the risk of sampling bias associated with motivation to take part in this study. It would appear that elderly Greek individuals willing to participate in the study are more motivated and possibly more curious of what a neuropsychological examination involves. Finally, an important limitation imposed upon the present study and other studies of MCI arises from lack of consensus regarding the clinical definition of MCI. Although our participants were selected according to the standard definition in the field (Petersen et al., 1999; Petersen, Stevens, et al., 2001), criteria for aMCI require mainly memory impairment. However, consistent with the findings in the literature (Constantinidou et al., 2014; Estevez-Gonzalez et al., 2003) and the results obtained by our aMCI patients on the brief neuropsychological battery (results not reported here), which was administered together with the RAVLT, these patients are often impaired in other cognitive domains. A more valid classification of this pre-dementia state may therefore be necessary in the future.

Finally, although this study focuses on the unique clinical utility of the RAVLT in diagnosis (e.g., the RAVLT is sensitive to early decline in aMCI and AD patients), there is a large possibility
that several of the other measures in our neuropsychological battery would have similar predictive power (e.g., perhaps RAVLT is related to global decline, so many measures would do as well).

Despite these potential limitations to the generalizability of the present normative data, the study provides a reference point for the neuropsychological assessment of verbal learning and episodic memory in elderly Greek adults over the age of 60 and extends the available Greek norms to elderly patients 89 years of age. Moreover, the study is based on culture- and language-specific learning lists that prevent Greek clinicians from having to inappropriately rely on English-word lists and norms for English-speaking or other populations in the elderly.

Even more importantly for clinical practice, the study provides evidence that the Greek version of the RAVLT adequately distinguishes the performance of predementia or amnestic MCI patients from that of cognitively healthy elderly, underlining its usefulness in the differential diagnoses of early dementia of the Alzheimer’s type. Moreover, the test is able to differentiate among major episodic memory decline of AD patients and the more minor verbal episodic memory impairments of aMCI patients.

Future research is required in order to collect normative data for the Greek nondemented oldest old (90+) and in order to establish the diagnostic utility of the Greek learning lists in differentiating among Alzheimer’s dementia and other dementia subtypes for, for example, behavioral variant frontotemporal dementia.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**References**


Lee, M. C. (2003). Normative data: Neuropsychological measures for Hong Kong Chinese. Hong Kong: The University of Hong-Kong, Neuropsychology Laboratory.


